Accommodation in Instant Messaging

A Senior Honors Project Presented to the Faculty of the
Department of Linguistics, University
Of Hawai'i at Mānoa

In Partial Fulfillment of the Requirements
For Bachelor of Arts with Honors

By Ivana Matson
March 24th, 2017

Committee:
Amy Schafer, Mentor
Katie Drager

Acknowledgements

I would like to thank everyone who has supported me through the writing of this thesis. Most importantly I must acknowledge my thesis advisors for their expertise and freely given time. I also acknowledge anyone else who has been a source of support along my academic journey.

Abstract

Differing levels of formality are salient to linguists and laypeople alike (Coupland 2014) - including in computer-mediated communication (CMC), which has not been extensively researched. As formality is so salient, it would not be surprising if users of CMC react differently to conversation partners who use different levels of formality, for instance by accommodating to their conversation partner's formality level (Beebe & Giles 1984). In this study, sociolinguistic interviews were carried out over instant messaging - a synchronous, one-on-one form of CMC. To investigate accommodation, these interviews were carried out in two conditions: one where formal features were used, and one where informal features were used. During the last part of the interviews in both conditions, the conditions were switched. Standard capitalization is a common marker of formality in CMC (Lahti & Laippala 2014, also shown to be perceived as true by frequent CMC users in a pre-study); due to this, one of the main differences between these conditions was that the interviewer used standard capitalization in the formal condition, but no capitalization at all in the informal condition. The results indicate trends in the expected direction, as well as substantial variation across participants.

Keywords: accommodation, instant messaging, computer-mediated communication

Table of Contents

List of Figures and Tables

1. Introduction

Speech accommodation is a commonly-attested phenomenon in sociolinguistics (e.g., cite); in contrast, instant messaging is a rarely-studied medium. This study seeks to bring the two together by investigating whether and how accommodation occurs over instant messaging, thus providing an example of a sociolinguistic process extending into a new medium, and showing that instant messaging operates in some ways much like speech, despite being written. Unlike spoken and signed languages, written forms of spoken languages are no one's native medium. For this reason, accommodation existing within a written medium would show flexibility on the part of those producing it, and perhaps reveal something about how speakers acclimate to new mediums.

The current study uses a variationist sociolinguistic analysis to investigate accommodation in instant messaging. Specifically, the study examines participants' convergence on different levels of formality, in terms of the text-specific features of capitalization and periods, used by the interlocutor (who was also the experimenter). Convergence to such text-specific features would show participants' adaptation to the medium.

2. Background

2.1 Speech Accommodation

Accommodation theory, as outlined in Beebe and Giles' 1984 paper, posits that one major reason for a given person's choice of speech style or register is adaptation to that of the person they are talking to. This adaptation could be towards the style of their conversation partner (convergence), or away from it (divergence), depending on factors such as whether the speaker wants to be associated with or seen as similar to their conversation partner or not.

Importantly, this accommodation is based on how the speaker *perceives* their conversation partner to be speaking. If a speaker seeks, for instance, to converge, this convergence may not be toward the exact style in use by their conversation partner, but rather the one the speaker thinks is being used (Beebe & Giles 1984).

This speaks to the degree to which people can influence each other's speech, which affects the process of language change. Linguistic features can spread via accommodation (Pardo 2006, Yu et al 2013). But as accommodation involves linguistic features that are already parts of the registers available to each interlocutor (Coupland 2014), it also reveals aspects of the social complexities of language use, for instance in terms of which feature is considered appropriate for what context. Accommodation also depends on social factors such as liking an interlocutor more (Yu et al 2013), biases towards an interlocutor's demographic (Babel 2010), and perceived similarity to the interlocutor (Weatherholtz et al 2014).

One of the most studied possibilities for accommodation is accommodation to an interlocutor's formality level. This is sometimes also associated with socially prestigious accents (Giles et al 1973). This subject of research is accessible because it is often already known what is considered "formal" in a speech community. It is also possible to manipulate perceived formality levels in ways other than changing speech styles, i.e. by interviewers dressing a certain way (Giles et al 1973) or being a particular ethnicity (Rickford & McNair-Knox 1994). In this study, formality was chosen as a framework primarily due to "formality" within written language already being well-known.

Many studies showing accommodation — including the present study— use interviews as a way of eliciting data. This is a convenient and effective tactic, as the interviewer can make decisions to speak a certain way, and recordings of the interviewee can be analyzed to check if

something in their own speech matched that of the interviewer, especially as in comparison with another interviewer, or with another type of data elicitation, like a wordlist. However, there are a few drawbacks with this method. One is that no interview is ever truly a natural context (Wolfson 1976). As people go about their daily lives, sitting down for an interview is simply not a normal activity. This is perhaps especially true for an "informal" interview of the type often used in sociolinguistics, where it hardly seems like specific questions are asked and participants may be confused as to the "point" of the interview, as Wolfson points out. Thus, there is always some influence on the data by the interview format itself, a reality made explicit in the Observer's Paradox, whereby the presence of an observer collecting data itself always influences the data (Labov 1972, p.113). In this study, traits of how the interviewer uses language are specifically used as the difference between experimental conditions. In other words, the presupposition that an observer influences any given speaker is here being used as the context for the experiment. Which way of speaking is "natural" is not at issue; whether the participants differ from each other depending on the way the interviewer speaks is.

One other way that has been used to "get around" the Observer's Paradox, and another factor to consider in interviews is the tendency for speakers to use more informal or less standard forms as interviews progress (Coupland 2014). This may be due to an increased sense of familiarity and comfort with the interview or interviewer "getting used to it," in essence. This increased familiarity may lead to more casual speech, often presumed to be more "normal" and less influenced by the observer (Coupland 2014). This can occur regardless of other factors of how a given interview is conducted and, foreshadowing the results, turned out to be an important factor in the current study.

2.2. Instant Messaging

Instant messaging (IM) is commonly used by young people in the United States, and has been since the early 2000s (Flanagin 2005, Grinter & Palen 2002). In the United States, use of IM predates use of texting, though in much of Europe, texting caught on faster (Ling & Baron 2007). Instant messaging is synchronous, one-on-one text-based communication via one of several computer programs designed for the purpose and is distinct from texting via SMS on a phone. IM, unlike SMS, takes place on a computer, via specific programs which may also mediate details like the formatting of italics and bold text. Both IM and SMS can be synchronous in the sense that a conversation partner receives one's message as soon as it is sent. However, unlike in speech, this message can be easily ignored, especially if a user is not actively checking their device. Also unlike in speech or signed language, where one can perceive a conversation partner's individual phonemes and words as produced in real time, messages are only sent once complete. Though the use of IM has diminished somewhat since the proliferation of social media (Quan-Haase & Young 2010), it is still a common form of computer-mediated communication (CMC).

IM is a medium that is used for purposes similar to those of speech. Though it is sometimes used to meet new people, it is much more commonly used to stay in contact with people the user already knows (Bryant et al 2006). However, it is not necessarily the case that who a user instant messages most frequently corresponds directly to who a user is closest to offline (Bryant et al 2006). IM is used primarily to maintain and develop relationships, and is used for this purpose more than social media (Quan-Haase & Young 2010). Ten years ago, it was outpacing e-mail as a relationship development tool (Flanagin 2005). IM is commonly used

in workplace settings (Voida et al 2002) as well as among friends, especially in young age groups (Lewis & Fabos 2005). Users of IM typically prefer face-to-face conversations when available, but will use IM as a second option, particularly in situations where it may be difficult to organize travel to meet in person (Lewis & Fabos 2005). IM does not seem to supplant existing offline friendships, but rather has been shown to increase their quality (Valkenburg & Peter 2009). As a method of communication used largely to develop relationships, much like speech often is, users of IM may develop tools within the medium to accomplish many of the same social functions that speech accomplishes.

One reason why IM is an interesting medium to study accommodation in is that it mixes styles and conventions of speech and writing. It is a written format with the speech-like characteristics of being synchronous and informal – but it can also be recorded, referred to, and even directly quoted to others (Lewis & Fabos 2005). As Lewis & Fabos put it in their 2005 paper, "IM users have to be good at sounding as though they are speaking in written texts." Instant messaging also has some practical considerations that are different from those of speech. In order to be seen as participating in a social interaction, one must actively input text, unlike in face-to-face contexts, where merely being present carries some degree of social participation (Lewis & Fabos 2005.) But IM also does not have to be immediately attended to – a conversation *can* carry despite pauses as one or both interlocutors become distracted by some other task (Voida et al 2002). This creates an unusual situation compared to speech.

The style of text produced in IM takes features both from speech and writing. In IM, innovative and archaic forms often exist side by side (Tagliamonte & Denis 2008, p.18). Prescriptive written standards play a role, which can become apparent when users of IM feel a need to edit or point out typos they make in order to make it clear that they understand they made

a mistake (Voida et al 2002). However, such prescriptive norms can also be purposefully disused in a meaning-making way, as in the distinction between "you" and "u" (Tagliamonte & Denis 2008), where the former is a formal and standard variant, and the latter's nonstandard nature can be used to convey a sense of friendliness or playfulness.

Conflicts between expectations derived from speech and writing – for instance the spoken language expectation of constant focus on the conversation, and the writing expectation of the conversation being non-synchronous – can create tension in IM (Voida et al 2002). These tensions can be dealt with using unique features of the medium. For instance, a long thought can be sent in multiple messages separated by line breaks to hold the conversational floor (Lewis & Fabos 2005).

As IM is a one-on-one, private form of CMC, it can be difficult to study, and some scholars have lamented that it has not been very well studied linguistically (Lewis & Fabos 2005, Tagliamonte & Denis 2008). In particular, there is very little research that has been done on IM from a quantitative perspective. As quantitative work can allow for generalizations about what an overall speech community is doing with their speech – or in this case text – it can be an important part of the overall research of a medium or speech community.

2.3 Predictions

Since accommodation is a very common process in speech, and instant messaging is in many ways speech-like, it was expected that evidence of accommodation would be found in IM. Specifically, it was expected that instant messengers would accommodate, on average, to an interviewer's use or nonuse of text-specific metalinguistic forms that are correlated with formality: standard capitalization and periods. This would be in line with some studied cases of

accommodation in speech that is linked with perceived formality, such as the use of -[ɪŋ] versus –[ɪn] (Fischer 1958, Rickford & McNair-Knox 1994).

The fact that the features analyzed in this study for accommodation are text-specific implies that any accommodation that occurs is not simply imported from speech, but rather is a novel mode of accommodation to match the available features of a new medium.

3. Methods

3.1. Procedure

Interviews were conducted using the instant messaging function on Skype. This program was chosen because of its ubiquity among instant messaging programs. Regardless of what programs participants most commonly messaged on, they consistently had access to Skype. Additionally, many participants stated that they used Skype most commonly out of all instant messaging programs. The entire interview process, from consent form to debriefing, took place over instant messaging. At no point was any participant's voice heard or face seen by the researcher.

Interviews were conducted under two conditions: "Formal" and "Informal." These conditions differed only in the interviewer's use of capitalization and punctuation. The same questions were used for both conditions. In the "Formal" condition, the interviewer consistently capitalized the beginnings of sentences, names, and the personal pronoun "I"; and sentences were ended with periods. In the "Informal" condition, no capital letters or periods were used, and commas were dropped where they were not absolutely necessary to disambiguate the meaning of a sentence. When organizing the beginning of an interview – for instance, asking a participant if they were ready to start – an effort was made to remain stylistically in the condition the interview would ultimately be in. Unfortunately, this was not done with complete

consistency.

After the pre-interview period, interviews began with a set of demographic questions about gender, age, ethnicity, and language background. Starting with these demographic questions, conditions were consistently applied, and remained consistently applied for the duration of the interview. Following demographics, questions that were described to participants as "general life" questions were asked. The purpose of these questions was to elicit as much data as possible, and hopefully inspire participants to type freely and naturally. These "general life" questions included "What would you do if someone gave you a billion dollars?" and "What was the worst attempt at cooking you ever made?" Following the "general life" section, participants were asked questions about the internet and language. These questions were designed to potentially provide qualitative insight into the subject of instant messaging, starting with the broad question "Have you noticed anything about the way people talk on the internet/in text messages compared to how they talk in real life?" and ending with the specific question "Do you think there's a way to talk over the internet, text, or IM formally? How would that be different from texting or IMing casually?"

After the internet and language section was the switch section, where participants were asked what was described to them as "wrap-up" questions. These questions were asked in the style of the opposite condition to what the rest of the interview had been conducted in. Therefore, for the "Formal" condition, the wrap-up questions were asked without the use of capitalization or periods, and for the "Informal" condition, the wrap-up questions were asked with standard capitalization and periods. The purpose of this was to provide a within-participant comparison between styles. The questions asked also elicited useful information regarding participants' use of computer-mediated communication in general. Questions in the wrap-up

section included “What social media sites do you tend to use?” and “Does autocorrect affect your IMing?” The questions in the switch section of the interview were of a type that could have caused participants to become more aware of their language use or self-conscious about it. This possibility is analyzed further in the discussion section of this paper.

At the end of this section, the interview was concluded. Participants were given written information about the purpose of the interview, and a small drawing containing the words “thank you” that the interviewer had made on a computer program. The style used at the end of the interview – the opposite from the primary condition – persisted in many cases past the end of the interview proper, and into any non-coded conversation that occurred regarding the purpose of the interview.

3.2. Participants

Participants were recruited by word-of-mouth, through flyers and social media postings on Facebook and Tumblr, and through the Language Beyond the Classroom program, via which undergraduates in introductory linguistics classes obtained class credit for participation. Recruitment messages included information that the interviews would be over instant messaging, and would be about the internet and language. No incentive, monetary or otherwise, was given to the participants recruited by word-of-mouth, and the Language Beyond the Classroom participants were given no incentive other than class credit.

The criteria for participation were: (1) they frequently instant messaged in English, (2) they were fluent in English, and (3) the interview was conducted while participants were at their computers (rather than using their phones). Of the 35 people who took part in the study, 19 participants were removed because they did not fit the criteria for participation.

Of the remaining 16 participants, 10 identified as female, five identified as non-binary and one identified as male. Eleven were from the United States, two were from the United Kingdom, one was from Canada, one was from Australia, and one was from Brazil. Ten identified as white, one identified as East Indian, one identified as Filipino, one identified as Malaysian, one identified as African-American, and two identified as mixed. 15 of the 16 participants were recruited via word of mouth, while one participated for extra credit in a linguistics class.

The excluded participants, when asked in the follow-up questions at the end of the interview about what instant messaging programs they used, said they did not use any instant messaging, said they used it rarely, or listed programs which were not instant messaging programs, but rather texting programs. The relevant difference here was whether the programs could be used on a computer and therefore without the automatic influence of autocorrection (e.g. Skype, Discord, or Facebook Messenger) or if they could only be used on a phone (e.g. Line or Textra). Their lack of use of instant messaging calls into question some previous assertions, for instance Tagliamonte & Denis 2008, that instant messaging is a favored medium among members of the younger generation. Most of these excluded participants frequently texted, suggesting it is possible that instant messaging is becoming less common as texting becomes more prevalent.

3.3. Coding methods

Data was coded only for the general life and switch sections of the interview. This was because the general life section was intended to have the most free and natural interaction of all the parts of the interview, and the switch section was the one that allowed for within-participant comparison of the conditions.

Three dependent variables were coded: the participants' use of standard capitalization, the participants' use of message-final periods, and the presence of message-medial periods. These variables were chosen because, after some assessment of dependent variables after the conducting of the first five interviews, these three were frequent in the data. Other measures that were considered, such as use of emoticons or unusual capitalization, were discarded because they simply occurred too infrequently. Message-medial periods, occurring when a single message contains more than one sentence, were considered as another measure to be coded. However, they were decided against due to the difficulty of determining possible locations where they might occur.

Use of standard capitalization and use of message-final periods were both coded as straightforward ratios. Sometimes other things were going on in a sentence that conflicted with a simple binary measure of "capitalized or not" or "period or not," and resulted in some cases that could not be coded. For instance, sentences ending in an exclamation mark were not coded regarding presence of a period. Similarly, if a sentence began with a multiple-letter word in all caps, it was not coded regarding presence of capitalization. This is because it is worth differentiating a sentence that ends in no punctuation from a sentence that ends in an exclamation point, and a sentence that begins in all lowercase from a sentence that begins in all uppercase. Outside of cases like these, every beginning of a sentence, every name, and every personal pronoun "I" were coded in a spreadsheet as either capitalized or not. Similarly, every sentence that could be was coded as either ending in a period or not.

4. Results

In this section, trends for capitalization are shown first and then the results for the use of periods is presented. For each variable, the raw data are presented first, followed by statistical models used to test the significance of any trends that were evident in the raw data.

For the analysis, interviews were divided into four sections. The first three sections are even thirds of the general section of the interview. They were split this way in order to test whether changes occurred over time during the general part of the interview. The size of the sections (i.e. the choice of thirds over halves or fourths) was chosen due to thirds of the general section being roughly the same length as the switch section.

4.1. Capitalization

Figure 1 shows the capitalization data for each subsection of the interview for each condition – formal on the left, informal on the right. The first three boxes in each graph are for the general section of the interview, divided into even thirds. The fourth is for the switch section, where the interviewer used the opposite typing style to that of what they used in the rest of the interview (i.e., an informal style in the formal condition and a formal style in the informal condition). The line within each box is the median of the participant capitalization proportions for that section. The box represents the interquartile range for those proportions, comprising the middle 50% of the data. The vertical lines outside the boxes each extend to the data point closest to the box which is still within 1.5 times the interquartile range.

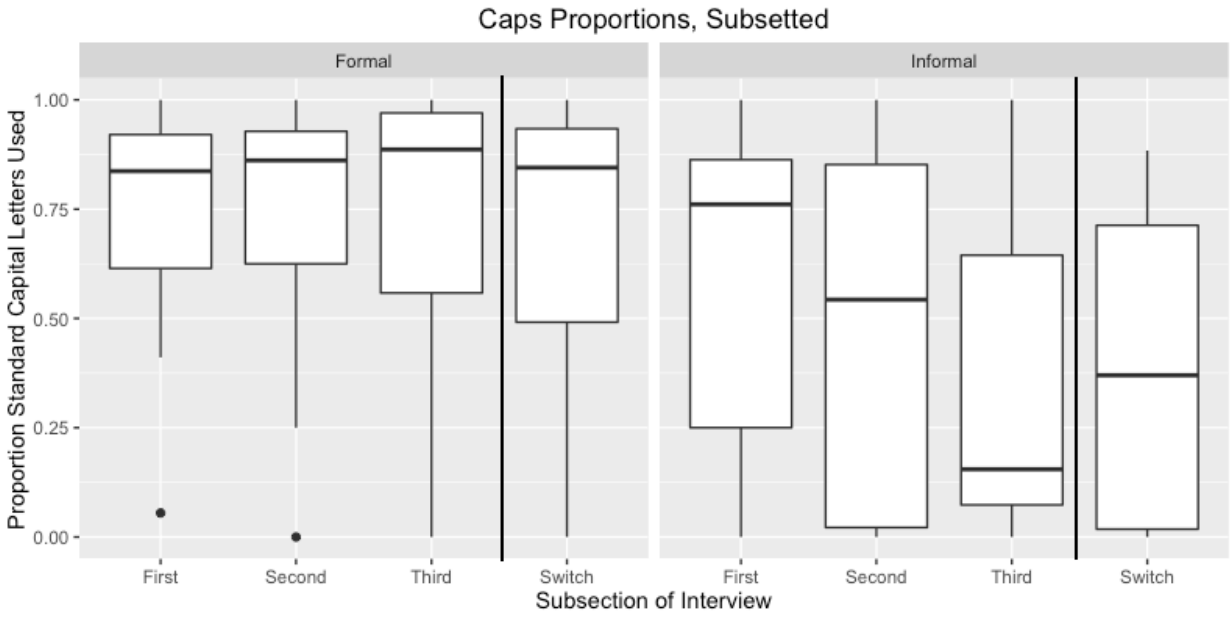

Figure 1: Subsetted capitalization data

Comparing the two plots in Figure 1, there are differences in medians and interquartile ranges across all four sections of the interviews. The interquartile ranges are weighted more toward frequent capitalization in the formal condition than in the informal condition. This is consistent with the study's predictions.

In the formal condition, in the general section, there was, overall, a very slight increase over time in the use of capital letters for some participants. However, some participants shifted in the opposite direction. In the switch section, when the interviewer was not using capital letters, participants' use of capital letters decreased to the lowest median of any point in the data. In the informal condition, use of capital letters decreased over time, as evidenced by the medians. However, there was a large amount of variation throughout, as evidenced by the large size of the interquartile ranges. Once the experimenter started using the more formal style (i.e., the switch) participants used a larger proportion of capital letters than they had previously. While the switch did not induce more capital letters than used at the beginning of the general section, the change

in the direction of the shift in previous sections is compelling. In both conditions, variation was high, although it was higher in the informal condition than the formal condition.

It is also useful to look at the difference between the general section overall and the switch section, without any subsetting into thirds. The general section, without subsetting, was about three times as long as the switch section, so the two boxes in each graph are formed from different amounts of data. Note that Figure 2 shows relatively little difference between the general section and switch section in either condition, though there is somewhat larger variation in the switch section of the formal condition than the general section. Once again, the formal condition is shown on the left and the informal on the right; within each graph, the box on the left shows the general section, and the box on the right shows the switch section.

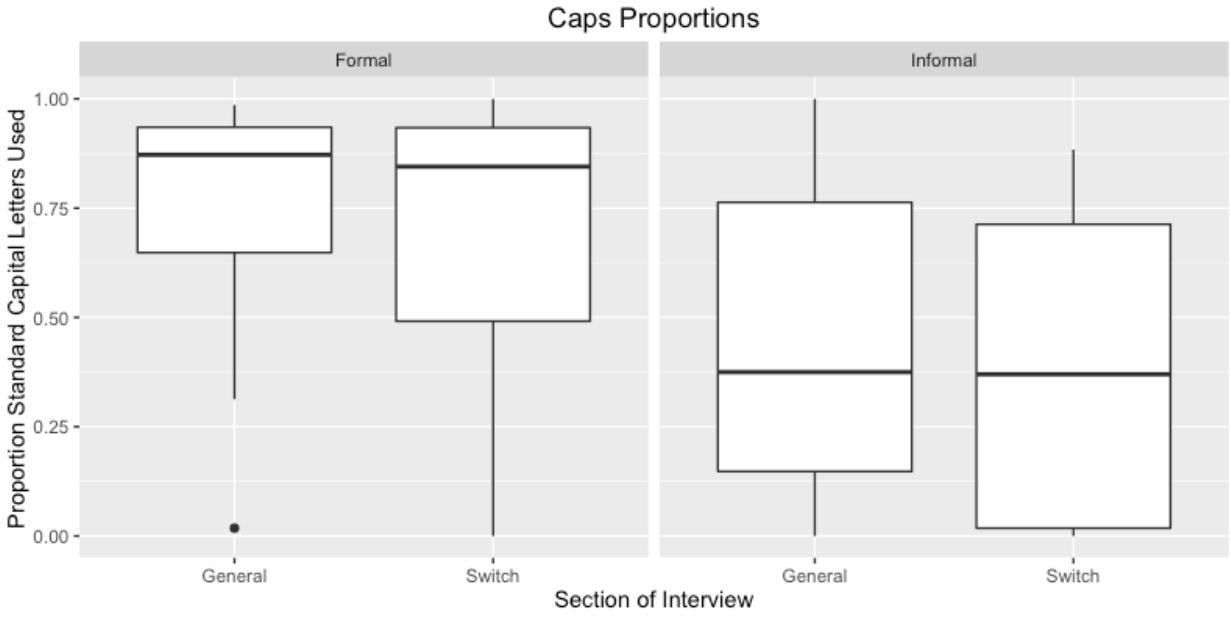

Figure 2: Non-subsetted capitalization data

To test the statistical significance of the interaction between condition and block, a logistic regression model, with slopes, was fit to the data using R. The model included condition (centered) and block (treatment coded) as fixed effects, and the maximal random effects structure

permitted by the design. Likelihood ratio tests verified that including by-participant slopes in the random effect structure improved the model. Section of the interview was also tested but was not included in the final models since it did not reach significance. The output of the model is shown in Table 1. A higher estimated coefficient indicates greater use of capitalization.

	Estimate	Std. Error	Z value	Pr(>|z|)	
(Intercept)	0.3842	0.7151	0.537	0.5911	
ConditionCtr	-1.7486	1.4168	-1.234	0.2171	
BlockSwitch	-0.7453	0.372	-2.003	0.0451	*
ConditionCtr:BlockSwitch	-0.3401	0.7456	-0.456	0.6483	

Table 1: Output of the overall capitalization model

The only factor to reach significance in the model was a main effect of block: participants capitalized less in the Switch section of the interview than in the General section. While the interaction between block and condition failed to reach significance ($p>.6$), the significant effect of block is likely carried by the large decrease in use and capital letters from the General section to the Switch section in the formal condition. This is in line with one of the predictions outlined earlier, which is that the interviewer's switch from a formal style to an informal one would lead to participants themselves becoming more informal.

To show that the effect of block is carried by the formal condition, separate models were fit for each condition. The results from the formal condition are shown in Table 3, and results for the informal condition are shown in Table 2. As shown in Table 3, the effect of the switch is not significant in the informal condition ($p>.2$). However, it reached significance in the model fit only to data from the formal condition ($p<.05$).

	Estimate	Std. Error	Z value	Pr(>|z|)	
(Intercept)	1.3565	0.8555	1.586	0.1128	
BlockSwitch	-0.6667	0.3051	-2.185	0.0289	*

Table 2: Output of the capitalization model fit to data from the formal condition

	Estimate	Std. Error	Z value	Pr(>|z|)	
(Intercept)	-0.3624	1.2033	-0.301	0.763	
BlockSwitch	-0.8791	0.7196	-1.222	0.222	

Table 3: Output of the capitalization model fit to data from the informal condition

Although the interaction was not significant in the overall model, the effect of block reached significance in the model fit only to the informal condition. The lack of statistical significance for the interaction in the overall model may be due to the relatively small number of participants and the large degree of variation across them. This large degree of variation is shown more clearly in Figure 3.

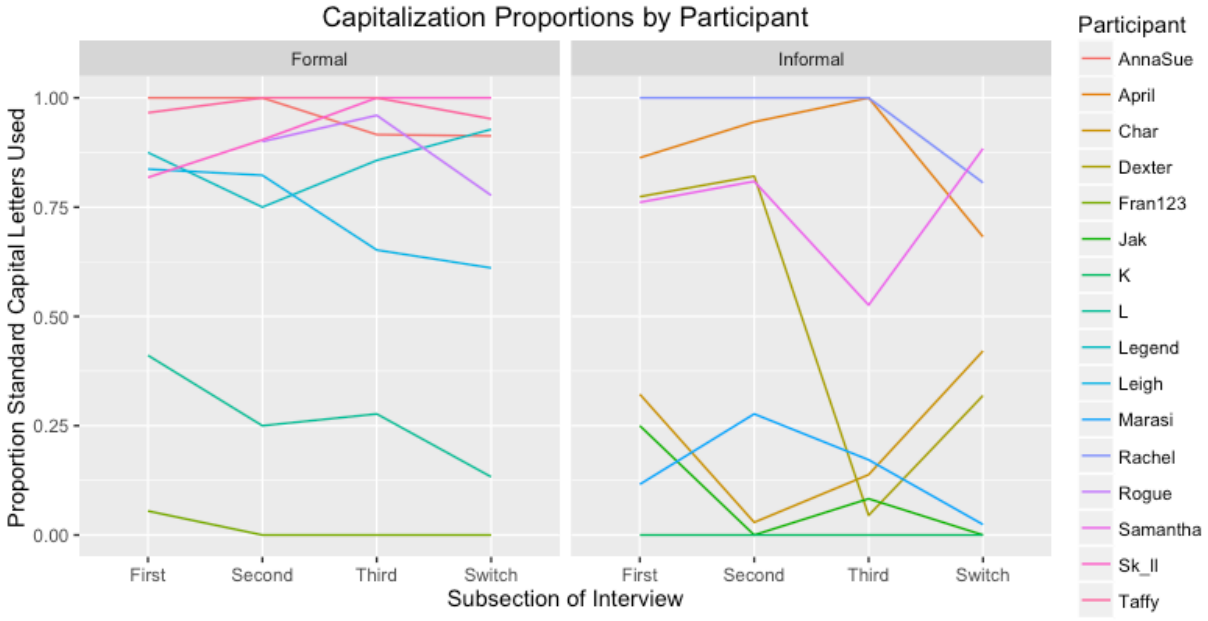

Figure 3: Capitalization over time by participant

Figure 3 shows line graphs of each participant's capitalization proportions over the subsets of the interview, divided by condition. Immediately evident in these graphs is that participants did not in general behave similarly to each other. Also notable is that few individual participants capitalized near 50% of the time. Many were closer either to 100% or 0% than 50%. The variability was somewhat higher in the informal condition than in the formal condition.

4.2 Periods

Figures 4 and 5 show the proportion of periods across the different subsections of the interview (in the first three boxes) and after the switch (in the fourth box). Data from the formal condition is on the left, and the informal condition is on the right. Figure 4 divides the general section into subsets in the same way as Figure 1; Figure 5 collapses the general section into a single box on each boxplot.

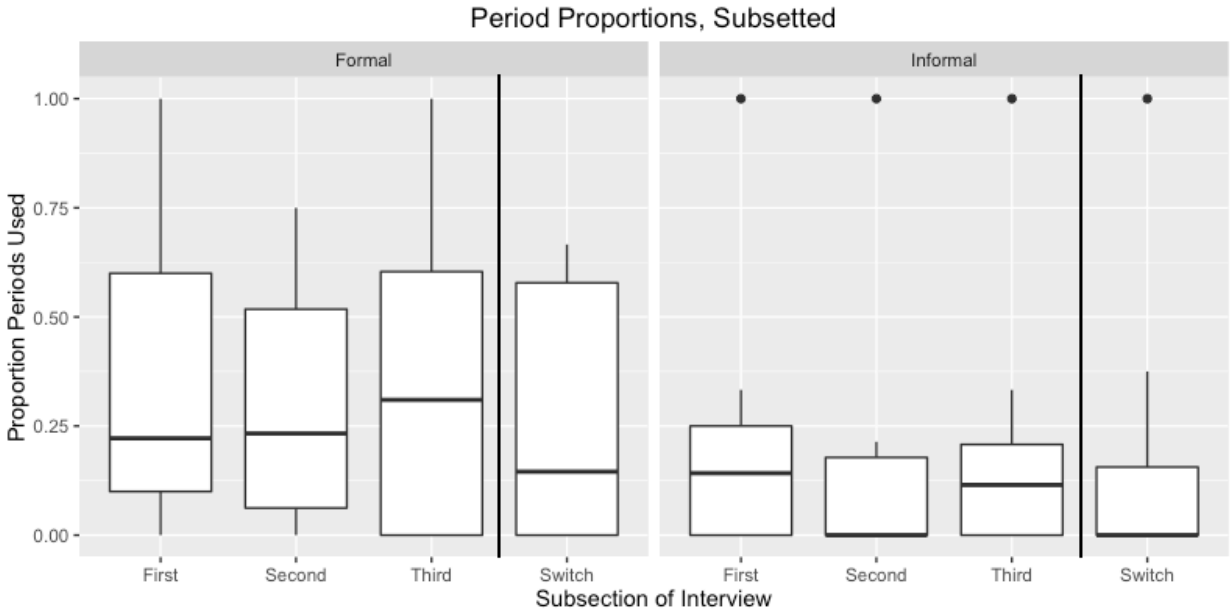

Figure 4: Subsetted period data

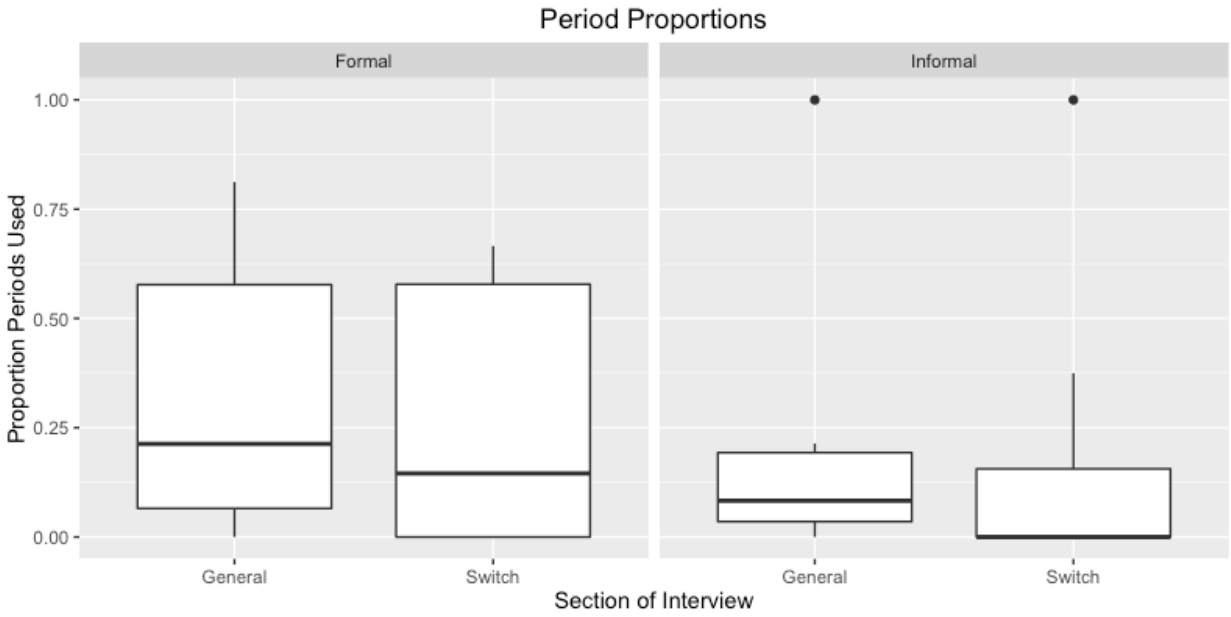

Figure 5: Non-subsetted period data

The conditions look notably different from each other in both Figure 4 and Figure 5. However, the medians between conditions are very close to each other. It is the interquartile ranges which are very different between the formal and informal conditions: they are much greater in the formal than in the informal. This is shown further in Figure 6, a line graph similar to Figure 3.

In the formal condition, Figure 4 suggests that there appears to be a slight increase in the use of periods over time (or at least between the second and third subsection), up until the switch, when the use of periods declined. In the informal condition, there was no notable trend. One potential reason for this lack of trend is a floor effect: it is impossible to use a lower percentage of periods (or of anything) than zero. Since participants in the informal condition tended to use an extremely low percentage of periods throughout the interview, the fact that there is a limit to how low a percentage can be might make it impossible to see much impact from the interviewer's style. This potential floor effect in the informal condition may also explain the

difference in interquartile ranges between conditions. Several statistical models were run, but no effect or interaction reached statistical significance.

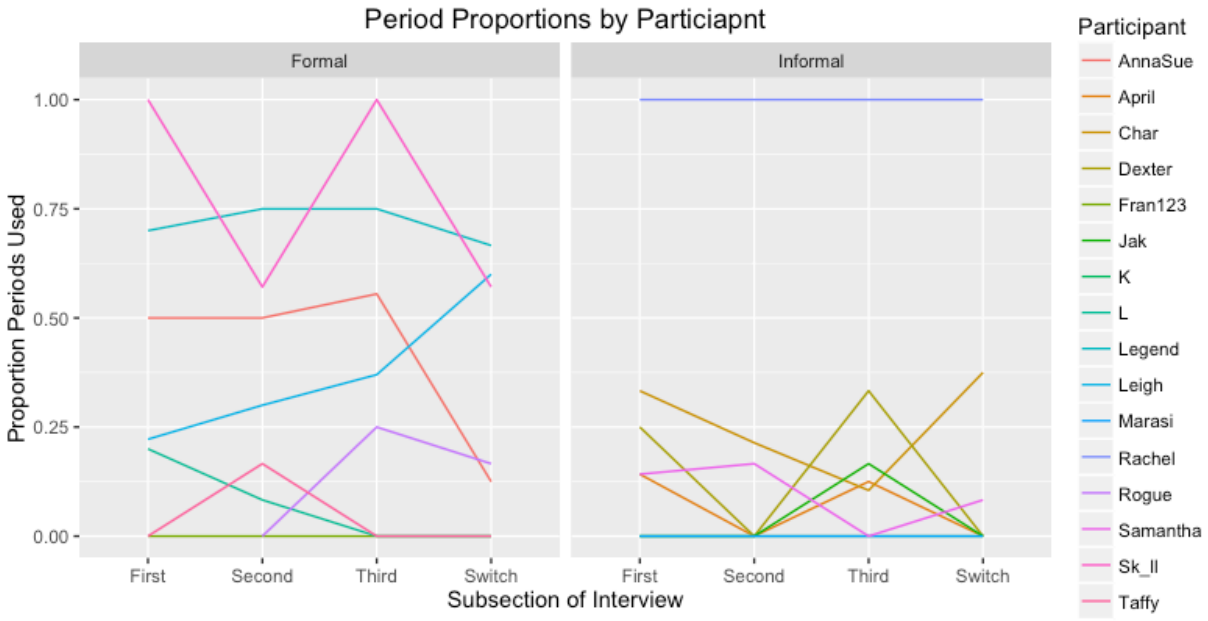

Figure 6: Period use over time by participant

Figure 6 clearly shows the much greater degree of variability in the formal condition than that of the informal condition. Almost every individual participant in the informal condition used periods a low proportion of the time. In the formal condition, there was a wide variety of proportions of period usage.

In general, difference in period use between blocks as well as between conditions was very low, as Figure 4 shows, and as with capital letters, variability was high. Additionally, there were far fewer data points per participant for use of periods than for use of capital letters: in some cases, as few as two or three data points per subsection. It is likely that only a strong trend could be visible with such a small amount of data, but we see little evidence of such a trend.

5. Discussion

5.1 Summary of findings

Capitalization usage, and to a smaller extent message-final period usage, differed numerically between conditions. Capital letters and message-final periods were used more often in the formal condition (with means of 71% capitalization and 32% period use in the general section) than the informal condition (with means of 45% capitalization and 20% period use in the general section). However, differences in capitalization and period use were not statistically significant. Similarly, boxplots show a decrease over time in the use of capital letters in the general block of the informal section, and a slight increase over time in the general block of the formal section, but this effect was also not significant.

The only significant effect found in the analysis was a decrease in the use of capital letters in the switch section of the formal condition. This suggests that the interviewer's switch to a more informal style at the end of the interview may have influenced participants, resulting in convergence on the experimenter's more informal style. It is possible that there was some degree of accommodation beyond this, for instance to the style used overall, as the trends were in the expected direction. The high variability and low numbers of participants in this data make it difficult to draw a large number of conclusions.

5.2. Potential influences on data

One of the most notable elements of this data is its wide variability. There is a large amount of variation within each condition. Further, when looking at the medians, the differences between conditions are relatively subtle and, as a result, most medians for subsections are within the interquartile ranges of most other conditions. The subtlety of the difference between conditions together with the high degree of variability in the data are responsible for the lack of significance for condition. Statistical significance implies relative certainty that the difference

between conditions is not due to random chance, but this certainty is hard to come by when the variation even within a single condition is very high.

This variability also in part speaks to a trait of the participant population, and possibly to IM as a medium. Some participants had very consistent styles, and either never or always used capital letters and/or periods. For instance, the participant with the pseudonym "Rachel" used message-final periods 100% of the time and capitalized nearly 100% of the time, but she was interviewed in the informal condition. Likewise, the participant "Taffy" in the formal condition capitalized almost as much as Rachel, but nearly never used message-final periods. "Fran123," interviewed in the formal condition, capitalized rarely and used no message-final periods; "K," interviewed in the formal condition, never used either capitals or message-final periods. Thus, stylistically extreme participants of both types – the always-formal and the always-informal – were found in both conditions. It is possible that the reason for their stylistic extremity was personal preference. Another possibility is that extreme levels of formality could have been due to the perceived formality of an interview context overall. Rachel, the most formal participant, noted that she was typing formally, while also describing the existence of other text communication contexts, including most instant messaging contexts, where she would not be formal. This is quoted in Example 1. Even in less extreme cases, it is possible that some participants in both conditions were pushed towards more formal speech by the very nature of it being an interview, as opposed to normal conversation. In other words, it is very possible the Observer's Paradox played an effect. In the opposite direction, many of the participants knew me from other contexts, so they may have typed more informally to me than they would have an unknown interviewer. Personal style, the interview context, and familiarity all could have affected participants' decisions about how to type regardless of condition. These are all

considerations that also apply to in-person, spoken interviews, and have a possibility to affect IM interview data as well.

(1)

[1/23/16, 7:19:30 PM] Ivana: in what ways are you formal when emailing teachers?
[1/23/16, 7:20:28 PM] Rachel: Proper punctuation and grammar. I hesitate to use the phrase but cannot think of a better one so, I use 'long words' and such.
[1/23/16, 7:20:39 PM] Rachel: Much like how I am speaking with you now, actually.

When comparing the general and switch sections of the interview, there was also another potential driving force behind differences between them. While the general section asked questions of the type common for sociolinguistic interviews – questions about personal experiences, designed to inspire participants to talk without much self-consciousness, the switch section coincided with a switch of topics, where participants were asked to explicitly speak about language on the internet. The fact that participants were thinking about their language use could have influenced them during this section in a number of ways. Additional consciousness about the way they were typing could have inspired them to type in ways they considered to be more prescriptively "correct," pushing them toward formality. It also could have inspired them to "show off" some of their styles of language use, potentially pushing them toward informality. A likely example of this occurs in Example 2, although it does not have to do with capital letters or periods. Here, the participant "Dexter" is saying that they like the way internet-based communication can include usage of "extra" – prescriptively unnecessary – punctuation to create tone. They say this information in a way that includes the exact type of "extra" punctuation they are referring to.

(2)

[2/8/16, 4:34:25 PM] Ivana: are there any things you like about the way people talk on the internet/in text messages?
[2/8/16, 4:36:59 PM] Dexter: i guess? i like when people use? extra punctuation to create tone??

It is also worth considering potential implications of which sections of the interview overall were *not* coded and analyzed. At the beginning of the interview, questions about demographics were asked, and this entire section was not coded. It is possible that this section could show participants typing in a way that was particularly unaffected by the way the interviewer did, due to less exposure time to the interviewer's typing style. It is possible participants in this section used a more formal typing style at the beginning than they did elsewhere due to not yet feeling comfortable, or that they used a more informal style due to feeling the interview had not "really started" yet. Similarly, any interaction that took place after a participant had consented for their interview to be stored in a database – the very last question of the interview – was not coded. This included the "debriefing" process, where it was described to participants what the purpose of the study was. It is possible participants used a more informal style in this part due to feeling, accurately, that the interview was "over".

Additionally, there was another factor driving a few instances of capitalization. The main reason participants were requested to type on their computers was to avoid the use of autocorrection algorithms. This was also a major reason why only participants who often did IM via their computers, rather than their phones, were included in the data, so that it could be guaranteed the participants were used to a medium in which autocorrection does not frequently occur. Unfortunately however, autocorrection happening infrequently on a computer is not the same as it never happening. Participants were asked at the end of the interview if autocorrection affected it, and a few – including those shown in Examples 3 and 4 – mentioned that it led to occasional unintended capitalizations. Though most participants were not affected by

autocorrection, and even those who were did not find the majority of their utterances to be autocorrected, this is still a notable methodological problem, and an external influence on the data.

(3)

[1/24/16, 12:47:41 PM] Ivana: Did [autocorrection] affect your IMing in this interview?
[1/24/16, 12:49:06 PM] Char: It capitalised a few sentence beginnings, but that wasn't really a problem.

(4)

[1/30/16, 4:31:22 PM] Ivana: Did it affect your IMing in this interview?
[1/30/16, 4:31:42 PM] Marasi: yeah a bit
[1/30/16, 4:31:53 PM] Ivana: In what ways?
[1/30/16, 4:32:34 PM] Marasi: most of the "i"s that are capitalized are a result of autocorrect
[1/30/16, 4:33:14 PM] Marasi: any proper nouns that are capitilzaed have been autocorrected

In addition to these issues of possible or near-certain influences on the data, it is also worth considering that there quite simply were not many participants. Though the initial goal had been for there to be thirty participants, many of those interviewed turned out to rarely use their computers for IMing, but instead use their phones. As has been discussed in the Methods section, these participants were not included due to the possibility that they would be used to autocorrection even when typing in a medium (i.e. computer-based instant messaging) that does not typically have autocorrection. One complication that could have arisen would be if these habitually expected capitalization to be added for them, and therefore did not capitalize even when they intended messages to be capitalized. Another possibility would be that if autocorrection regularly caused all their text-based communication to be capitalized, use or nonuse of capitalization would never have a chance to develop a social meaning for them. The small number of total participants, of course, means that any single participant had a fairly large

effect on the overall data in a given condition, which may have contributed to the lack of significance.

It is also worth considering that the participants were demographically similar, and this may have affected results. It is possible the results were pushed toward the norm of the participants' demographics, even if that norm was itself high variability, for instance if people of the participants' demographics tended to categorically either use formal or informal styles. Many participants frequented the same social media websites and even the same fan communities. Most were white. Though both females and non-binary people were well-represented, there was only one male. On this last point, it is worth noting that most participants, when asked during the internet and language section of the interview whether they thought people of different genders talked differently on the internet, said they did not think different genders spoke differently on the internet. However, it is possible that even the participants' thinking this was influenced by their gender. Participants were indeed quite variable in their use of both capital letters and periods, suggesting that additional demographic variability might be unlikely to make results yet more variable. However, it could be that members of a different demographic may have norms which these participants did not, or may have reacted differently to conditions.

5.3. Further possibilities

Formality, informality, and registers and style in general are, of course, a matter of more than just capitalization and message-final period use. Even in this research, message-medial periods were strongly considered for study, but ultimately were not analyzed. This was due partially to their relative rarity, but also due to the difficulty of coding them. One possible option

in future research might be to measure punctuation use overall. Common use of periods as well as semicolons and commas may in most cases point to formality, although even here there might also be other factors to consider, as sometimes punctuation can be used in distinctly nonstandard ways. Beyond this, there are other, more difficult variables to quantify in text that may mark formality, including word choice and sentence structure.

One major difficulty in the present study was finding participants who fit the target demographics, largely because few people frequently used IM on their computers, opting instead to use their phones. Though this was considered to be a problem due to the potential interference of autocorrection, it also seems to be the case that CMC use today is different than it was in the IM-centric times of Grinter & Palen's 2002 research, Flanagin's 2005 research, or even Taglimonte & Denis' 2008 research. Currently, in the United States, it appears that texting and use of phone-based IM programs are both more prevalent than computer-based IM, at least if the difficulties of recruitment in this study are any indication. Moving forward, it may be worthwhile to investigate variation within this phone-based context. Though autocorrection may make the specific measure of capitalization more difficult to analyze on a phone, it seems reasonable to assume that if phone-based text communication is as common of a medium as it seems to be, there is likely to be socially-meaningful linguistic variation.

6. Conclusion

The significance of the decrease in capital letter use from the general section to the switch section of the formal condition provides some evidence for the presence of accommodation in instant messaging. However, extremely high variability in participants' use of capital letters as well as periods means that accommodation shown in the data, even within

this one significant effect, is small. Larger numbers of participants would almost certainly be required to more clearly show effects of accommodation in instant messaging.

However, even this slight display of accommodation in a fairly recent, text-based medium is interesting. Accommodation is an extremely common, natural process in spoken language; for it to appear in a medium that is no one's first-learned medium implies that that medium has become somewhat natural to those who use it. IM is being used in very much the way speech is used, as social communication with many small variables that can be changed to show alignment with one's interlocutor. The fact that the variable which showed some accommodation in this study – capital letters – is a text-specific variable is interesting. This variation is text specific, meaning it cannot simply imported from speech. This implies that the social functions of speech underlying accommodation also exist in the context of IM. The expression of these functions is in some ways similar to speech (accommodation) and in some ways different (text-specific variables). It is possible that as more text-specific variables attain social meaning – whether through the lens of formality versus casualness or something else – these variables will also begin to display accommodation or other elsewhere-attested sociolinguistic phenomena. Capital letters as a socially meaningful variable may just be the start of more text-specific variables becoming socially meaningful; they also may simply be the most obvious of a larger number of socially meaningful variables which currently exist in the medium.

References

Babel, Molly. 2010. Dialect divergence and convergence in New Zealand English. *Language in Society* 39.437-56.

Beebe, Leslie M., and Howard Giles. 1984. Speech-accommodation theories: a discussion in terms of second-language acquisition. *International Journal of the Sociology of Language.* 1984:46.5-32.

Bryant, Alison J; Ashley Sanders-Jackson; and Amber M. K. Smallwood. 2006. IMing, text messaging, and adolescent social networks. *Journal of Computer-Mediated Communication.* 11.577-92.

Coupland, Nikolas. 2014. Social context, style, and identity in sociolinguistics. *Research Methods in Sociolinguistics: A Practical Guide,* ed. by Janet Holmes and Kirk Hazen. John Wiley and Sons, inc.

Fischer, John L. 1958. Social influences on the choice of a linguistic variant. *Word*. 14.47-56.

Flanagin, Andrew J. 2005. IM online: instant messaging use among college student. *Communication Research Reports*. 22:3.175-87.

Giles, Howard; Donald M. Taylor; and Richard Bourhis. 1973. Towards a theory of interpersonal accommodation through language: some Canadian data. *Language in Society*. 2:2.177-92.

Grinter, Rebecca E, and Leysia Palen. 2002. Instant messaging in teen life. Proceedings of the annual ACM conference on computer supported cooperative work. 21-30.

Labov, William. 1972. Some principles of linguistic methodology. *Language in Society*. 1.97-120.

Lewis, Cynthia, and Bettina Fabos. 2005. Instant messaging, literacies, and social identities. *Reading Research Quarterly*. 40:4.470-501.

Ling, Rich, and Naomi Baron. 2007. Text messaging and IM: linguistic comparison of American college data. *Journal of Language and Social Psychology*, 26:3.291-8.

Pardo, Jennifer S. 2006. On phonetic convergence during conversational interaction. *The Journal of the Acoustical Society of America.* 119.2382-93.

Rickford, John R., and Faye McNair-Knox. 1994. Addressee- and topic-influenced style shift: a quantitative sociolinguistic study. *Sociolinguistic perspectives on register,* ed. by Douglas Biber and Edward Finegan. 235-76

Tagliamonte, Sali A, and Derek Denis. 2008. Lingustic ruin? LOL! Instant messaging and teen

language. *American Speech* 83:1.3-34.

Quan-Haase, Anabel, and Alyson L. Young. 2010. Uses and gratifications of social media: a comparison of Facebook and instant messaging. *Bulletin of Science, Technology, & Society*. 30:5.350-61.

Valkenburg, Patti M., and Jochen Peter. 2009. The effects of instant messaging on the quality of adolescents' existing friendships: a longitudinal study. *Journal of Communication*. 59.79-97.

Voida, Amy; Wendy C. Newstetter, and Elizabeth D. Mynatt. 2002. When conventions collide: the tensions of instant messaging attributed. Proceedings of the SIGCHI conference on human factos in computing systems. 187-94.

Weatherholtz, Kodi, Kathryn Campbell-Kibler & Florian Jaeger. 2014. Socially mediated syntactic alignment. *Language Variation and Change* 26 (3), 387–420.

Wolfson, Nessa. 1976. Speech events and natural speech: some implications of sociolinguistic methodology. *Language in Society*, 5:2.189-209.

Yu, Alan, Carissa Abrego-Collier, and Morgan Sonderegger. 2013. Phonetic imitation from an individual-difference perspective: Subjective attitude, personality and "autistic" traits. *PLoS ONE*. 8(9):e74746.